The role of task similarity for ego depletion: A registered report

Primoceri, Pierpaolo ; Ramer, Nicolas ; Ullrich, Johannes ; Job, Veronika

Abstract: Ego depletion refers to decrements in self-control performance resulting from prior use of self-control. The ego depletion effect has received much research attention, but the more recent literature reports small or null effects. This registered report examined the moderating effect of task similarity on the ego depletion effect. We predicted a crossover interaction between type of primary and secondary task such that engaging in a demanding self-control task should lead to better performance in the secondary task when it is similar to the primary task (facilitation effect) but worse performance when it is dissimilar (ego depletion). In a preregistered pilot study, N = 80 participants first completed either a visual stop-signal task (SST) or a simple lexical categorization task. They proceeded with one of four tasks classified as increasingly dissimilar based on their underlying operations and executive functions: (1) auditory stop-signal task, (2) Stroop task, (3) Eriksen flanker task, and (4) unsolvable anagrams. Both the pilot study and a high-powered registered replication (N = 300) revealed the predicted interaction effect. However, evidence for facilitation from similar tasks was stronger than evidence for depletion from dissimilar tasks. Together, these findings highlight the important role of task similarity for the study of ego depletion and related phenomena.

DOI: https://doi.org/10.1016/j.jesp.2021.104133

Originally published at:
Primoceri, Pierpaolo; Ramer, Nicolas; Ullrich, Johannes; Job, Veronika (2021). The role of task similarity for ego depletion: A registered report. Journal of Experimental Social Psychology, 95:104133.
DOI: https://doi.org/10.1016/j.jesp.2021.104133

The role of task similarity for ego depletion: a registered report

Pierpaolo Primoceri[†,1], Nicolas Ramer[†,1,2], Johannes Ullrich[1], and & Veronika Job[2,3]

[1] University of Zurich

[2] Dresden University of Technology

[3] University of Vienna

This manuscript has been accepted for publication in Journal of Experimental Social Psychology, published by Elsevier. Please cite as follows:

Primoceri, P., Ramer, N., Ullrich, J., & Job, V. (in press). The role of task similarity for ego depletion: a registered report. *Journal of Experimental Social Psychology.*

Author Note

The authors declare no conflict of interest.

Funding: This work was supported in part by a Doc.CH grant from the Swiss National Science Foundation awarded to Pierpaolo Primoceri [P0ZHP1_162308].

Acknowledgments: We thank Caroline Surrey and two anonymous reviewers for their helpful comments on an earlier version of this article and Dino Semadeni, Yalini Ahrumukam, Annette Metzger, Jiyoung Ryu, Lea Hütt and Maria Pohl for their help in data collection.

† Pierpaolo Primoceri and Nicolas Ramer contributed equally to this work.

Correspondence concerning this article should be addressed to Pierpaolo Primoceri, Binzmühlestrasse 14 / Box 15, 8050 Zurich. E-mail: pierpaolo.primoceri@uzh.ch

Abstract

Ego depletion refers to decrements in self-control performance resulting from prior use of self-control. The ego depletion effect has received much research attention, but the more recent literature reports small or null effects. This registered report examined the moderating effect of task similarity on the ego depletion effect. We predicted a crossover interaction between type of primary and secondary task such that engaging in a demanding self-control task should lead to better performance in the secondary task when it is similar to the primary task (facilitation effect) but worse performance when it is dissimilar (ego depletion). In a preregistered pilot study, $N = 80$ participants first completed either a visual stop-signal task (SST) or a simple lexical categorization task. They proceeded with one of four tasks classified as increasingly dissimilar based on their underlying operations and executive functions: (1) auditory stop-signal task, (2) Stroop task, (3) Eriksen flanker task, and (4) unsolvable anagrams. Both the pilot study and a high-powered registered replication ($N = 300$) revealed the predicted interaction effect. However, evidence for facilitation from similar tasks was stronger than evidence for depletion from dissimilar tasks. Together, these findings highlight the important role of task similarity for the study of ego depletion and related phenomena.

Keywords: self-control, ego depletion, facilitation, dual-task paradigm, task similarity, inhibition

The role of task similarity for ego depletion: a registered report

Does exercising self-control in one task impair self-control in a second, unrelated task? The possibility of such ego depletion effects has occupied social psychologists since the first demonstration by Baumeister et al. (1998). Despite the intuitive validity and numerous empirical confirmations using a sequential-task paradigm (Hagger et al., 2010), the effect has recently been challenged on multiple grounds, most notably publication bias (Evan C. Carter & McCullough, 2013; Evan C. Carter & McCullough, 2014) and repeated failures of replication (Hagger et al., 2016; Lurquin & Miyake, 2017; Osgood, 2017; Xu et al., 2014). In addition, recent theoretical reviews have pointed out the lack of a clear operational definition and, more importantly, a theoretical and functional link between the respective definition and the selection of tasks used to study the effect (Friese et al., 2019; Lurquin & Miyake, 2017). Thus, research on ego depletion would benefit from a more systematic conceptual analysis of the relationship between the two tasks used in the sequential-task paradigm. We developed a taxonomy of self-control tasks that might be useful in determining the similarity between such tasks. More importantly, based on this taxonomy, we propose a new design to test the hypothesis that gradations of similarity would be related to ego depletion effect sizes.

We build on the recent notion that if two tasks require the use of the same cognitive function, performance may improve rather than decline through cognitive adaptation (Dang et al., 2014). This proposition was derived from conflict adaptation theories, which deal with the recruitment of control processes and the adaptation of the human cognitive system to specific task demands (see Braem et al., 2014, for a review). Conflict monitoring theory (Botvinick et al., 2001), for instance, suggests that when a person is engaged in a task requiring cognitive control, performance is supported by adjustments in perceptual selection, response biasing, and on-line maintenance of contextual information. In other words, the cognitive system adjusts itself to the current task-specific demands, becoming more autonomous as it adapts. An example of such adjustment in cognitive control is the finding that Stroop interference effects tend to be low if the frequency of incongruent trials is high (Lindsay & Jacoby, 1994; Logan & Zbrodoff, 1979). The theory posits that such cognitive control selection and adjustment processes are recruited by a conflict monitoring system governed by the anterior cingulate cortex (ACC). Higher ACC activation predicts, for example, more focused responding and thus better performance in the Stroop task (Botvinick et al., 2004).

Arguably, most of these findings have been observed and studied on a trial-by-trial basis so that one may legitimately wonder whether such adaptation effects carry over from one task to another in the dual-task paradigm. Recent work on proactive control suggests that this may indeed be possible (e.g., Braver, 2012; Egner, 2014; Verbruggen, 2016). The key insight is that the adjustment of cognitive control to longer timescales is cognitively abstracted and can be maintained to facilitate control in a different but similar task. There is also empirical support for such proactive control mechanisms across longer timescales. For example, Gonthier et al. (2016) showed that the aforementioned effects of the relative frequency of incongruent trials in the Stroop paradigm can be attributed to proactive control processes. Furthermore, recent evidence suggests that, when context-dependent task features remain constant, the activation of inhibitory control processes can also improve performance over an entire task (Samrani et al., 2018) or across variations of task blocks (Surrey et al., 2017; Torres-Quesada et al., 2013). Taken together, there is theoretical and empirical support for the assumption that cognitive adaptation might contribute to improved performance in the secondary task of the dual-task paradigm, especially when both tasks are similar.

Indeed, previous work by Dewitte and colleagues supports such effects in the context of ego depletion (Dang et al., 2013; Dewitte et al., 2009; Xiao et al., 2014). They suggest that cognitive adjustment to specific task demands may result in reduced flexibility with regard to new tasks that usually require different control processes. From this perspective, ego depletion effects may be the result of prior activation of task-specific control processes lingering on and interfering with the adaptation to the new task and its specific demands (Dang et al., 2013), akin to cognitive "switch costs" (Kiesel et al., 2010). On the other hand, if task demands and the required control processes remain similar from one task to another, prior adjustment should facilitate rather than hinder performance on a subsequent task. In other words, given that both tasks in a sequential-task paradigm require self-control, there should be a certain overlap in the required cognitive functions. Thus, on top of the process leading to ego depletion, there may be a process leading to better performance in the second task. This latter process should vary in strength depending on the similarity between the first task and the second task.

In line with this reasoning, across three studies, Dewitte et al. (2009) found ego depletion effects with dissimilar tasks and enhanced self-control performance with similar tasks. For example, in one study they showed that participants were more successful in completing a response reversal task after first completing a similar response reversal task rather than a thought suppression task (Study 1). Similarly, participants were more likely to choose a high self-control option in hypothetical scenarios, when two successive scenarios included the control of similar experience (impulsivity vs. impatience; Study 3). Further, as part of a different set of studies, participants were repeatedly found to perform worse on a secondary task that was dissimilar from the primary task and required flexible response adaptation. In contrast, if the secondary task was similar to the primary task and stable with regard to the required cognitive response, participants performed better (Wenzel et al., 2013, 2014).

Nevertheless, some studies have also reported depletion effects within the same or similar tasks, which may question the proposed idea. Perhaps most notable are studies from the field of mental fatigue, which have consistently found performance decrements in mentally demanding tasks (e.g., Boksem & Tops, 2008; Brewer et al., 2017; Kato et al., 2009; Lorist et al., 2005). However, while the exact relationship between mental fatigue and ego depletion remains an ongoing and unresolved issue (Pattyn et al., 2018), there are at least two important differences between the two frameworks. That is, in contrast to the bulk of ego depletion studies, mental fatigue studies generally employ tasks that are much longer in terms of task duration and much simpler in terms of the cognitive processes they require (Giboin & Wolff, 2019). While continuous performance in these tasks is inevitably linked to motivational impairment (e.g., Boksem et al., 2006; Boksem & Tops, 2008; Brewer et al., 2017; Herlambang et al., 2019), task characteristics do usually not allow for cognitive improvement. For example, the psychomotor vigilance task used by Brewer et al. (2017) required participants to simply react to the start of a digital counter as fast as possible. Thus, apart from sustained attention, the task did not involve any cognitive processes that would have allowed for notable improvement or facilitation through prior activation.

In sum, a number of previous results provide evidence of facilitation effects occurring between self-control tasks, which stands in contrast to the general ego depletion hypothesis. This effect seems to be particularly pronounced when the underlying cognitive processes associated with the tasks overlap, suggesting a moderating role of task similarity for ego depletion. Simply put, ego depletion may be more likely to occur between dissimilar rather than similar tasks (Chatzisarantis & Hagger, 2015; but see Lange, 2015), whereas facilitation effects may be more likely to occur between similar tasks, counteracting any potential ego depletion effect. In view of the fact that hardly any study has taken the potentially crucial role of task similarity into account when applying a sequential-task paradigm, more direct evidence of such an effect would substantially advance research on ego depletion by providing one potential explanation for the current inconsistencies in the field and further improving the paradigm. More specifically, it could explain why most original studies run before the replication crisis are the ones that most consistently found depletion effects (cf., Vadillo, 2019), as they generally employed quite different and (often behavioral) domain-independent tasks. In contrast, most recent replication studies, which found small or no effects, have predominantly used computerized tasks that are rather similar (or at least insufficiently dissimilar) with regard to the required underlying cognitive functions (e.g., Hagger et al., 2016; Lurquin & Miyake, 2017; Xu et al., 2014).

The primary goal of the current research was to extend the reported findings on the relationship between task similarity and the ego depletion effect by examining and taking into account the associated cognitive processes more directly. To this end, we first derived a taxonomy of self-control tasks that vary in similarity on the basis of their underlying cognitive processes. This should ultimately increase our a priori knowledge of when to expect larger, smaller, or even reversed ego depletion (i.e., facilitation) effects. We present preliminary findings from a pilot study that support our taxonomy and our related predictions. On this ground, the final aim of this registered report was a high-powered test of our hypothesis:

> ***Hypothesis:*** Ego depletion effect sizes should be linearly related to gradations of task similarity, with the strongest ego depletion effects occurring at low levels of similarity and the strongest facilitation effects occurring at high levels of similarity.

A taxonomy of self-control tasks

A first classification of self-control tasks has been proposed and investigated by Hagger et al. (2010) in their prominent meta-analysis. Importantly, the authors examined whether studies employing self-control tasks that were matched with regard to their suggested classification presented different results than studies with unmatched tasks. In line with a global resource account of self-control suggesting a domain independency for ego depletion, the analysis showed virtually the same effects for matched and unmatched tasks. It is important to note, however, that the suggested categorization, on which the matching of depleting and dependent tasks was based, is highly generic and based solely on task context and/or specific task features. For example, any cognitive inhibition task (e.g., Stroop and flanker task) was classified as "impulse control task," irrespective of their underlying functional processes (cf. Friedman & Miyake, 2004). Moreover, this same category further included tasks such as letter-crossing, food taste perception, resistance to tempting food, the handgrip task, or several persistence tasks. In short, these broad categories do not allow for a more fine-grained differentiation and analysis of self-control tasks that take into account specific task demands (e.g., inhibiting vs. initiating actions) or their underlying cognitive processes.

In order to fill this gap, we derived a taxonomy that infers task similarity from the level of transferability of executive functions between two tasks (Enriquez-Geppert et al., 2013). Accordingly, we distinguish the transfer between (1) identical tasks with different stimuli (modality transfer), (2) tasks targeting the same domain/function (near transfer), (3) tasks targeting different domains/functions (far transfer), and (4) tasks targeting different contexts (meta-cognitive transfer). In order to classify tasks as involving the same or different functions, we assume that self-control refers to the processes involved in "overcoming immediate reactions in favor of a longer-term goal (or task)" (Wilkowski & Robinson, 2016, p. 302). This requires the down-regulation of undesirable responses (inhibition) and/or the up-regulation of desirable responses (initiation/persistence). Assuming that these core functions reflect two distinct cognitive factors of self-control (Maloney et al., 2012; Ridder et al., 2012), transfer between inhibitory and initiatory self-control tasks can be classified as meta-cognitive. A smaller degree of dissimilarity should characterize tasks within the domain of inhibitory control. Previous research suggests that inhibitory control can be broken down into distinct processes (Friedman & Miyake, 2004): Prepotent response inhibition, resistance to distractor interference, and resistance to proactive interference. For the purpose of the current research, we focus on the first two processes (cf., Gärtner & Strobel, 2021). Friedman and Miyake (2004) identify the stop-signal task (SST) and the Stroop task (Stroop, 1935) as exemplary tasks requiring prepotent response inhibition. In constrast, the Eriksen flanker task (Eriksen & Eriksen,

1974) constitutes an exemplary task requiring resistance to distractor interference.

Taken together, we suggest that such self-control tasks can be categorized and matched with levels of transferability on the basis of their underlying cognitive functions. In line with the proposed taxonomy, Table 1 presents an exemplary classification of these common self-control tasks and their primary underlying functions discussed above. Accordingly, we chose a visual SST as our baseline task, which functionally involves the inhibition of a prepotent response. Following this rationale, each transfer level can be matched with an increasingly dissimilar task. In our case, this includes (1) an auditory SST, which differs only in terms of response modality (modality transfer), (2) a Stroop task, which differs only in terms of the task-related operation (near transfer), (3) a flanker task, which differs both in terms of the task-related operation and the underlying inhibitory function (far transfer), and (4) an unsolvable anagram task, in which the contribution of inhibitory processes should be minor relative to initiatory processes that boost persistence and problem-solving (Allom & Mullan, 2014; Heller et al., 2017).

Of course, our exemplary classification is based on the arbitrary choice to set the SST as a starting point. That is, one could just as well choose a visual color-word Stroop task as a baseline and map different tasks to the transfer levels of the taxonomy following the same principle (e.g., an auditory Stroop task for modality transfer, a visual SST for near transfer, etc.). Another plausible alteration might be to identify and map tasks in terms of distinct subfunctions of initiatory (rather than inhibitory) self-control. Consequently, the proposed taxonomy may be extended to the classification of other self-control tasks that share varying degrees of cognitive-functional similarity in future research (see, for example, Eisenberg et al., 2019 for a recent discussion of self-regulatory tasks). Nevertheless, we present a testable, theory-based classification of self-control tasks that is both supported by recent findings (e.g., Gärtner & Strobel, 2021) and highly relevant for the ego depletion literature.

In light of the enormous variety of self-control tasks that are still being used to study ego depletion (Lurquin & Miyake, 2017) the implication is that considerable work remains to be done to assess the similarity relations between these tasks. The goal of the present research was to obtain initial evidence for the utility of such a similarity taxonomy by varying the similarity of primary and secondary task in the dual-task paradigm according to the classification shown in Table 1. Our hypothesis maintained that these gradations of task similarity determine the direction (facilitation or depletion) and strength of the effect of initial self-control engagement.

Table 1

Exemplary Taxonomy of Self-Control Tasks in Order of Decreasing Similarity

Transfer level	Task	Operation	Inhibitory function
Baseline	Visual SST	Inhibit key press after visual stop-signal	Prepotent response inhibition
Modality transfer	Auditory SST	Inhibit key press after auditory stop-signal	Prepotent response inhibition
Near transfer	Stroop Task	Inhibit responding to semantic meaning	Prepotent response inhibition
Far transfer	Flanker Task	Ignore distracting stimuli	Resistance to distractor interference
Meta-cognitive transfer	Unsolvable Anagrams	Persist on difficult/unsolvable task	Resistance to impulse to skip

Note: SST = stop-signal task.

Pilot study

We conducted a pilot study to test the feasibility of our approach and obtain support for our reasoning. An a priori power analysis, hypotheses and planned analyses were preregistered (https://osf.io/7wuhp/) and closely followed. Minor deviations in handling response times are reported in the supplemental documents, available with all materials, data, and analysis scripts at https://osf.io/b9yek/. We report all measures, manipulations, and exclusions for the study.

Participants and Design

We recruited students from a Swiss university who received either course credit or CHF 15 in exchange for their participation. Participants were eligible if they were between 18 and 35 years old as well as fluent in the German language. They were tested in group sessions with up to four participants per session. All tasks and control measures were presented on a laptop running the experiment software OpenSesame (version 3.1.6 *Jazzy James*, Mathôt et al., 2012). All responses and response times (RT) were logged and measured with millisecond precision using a labeled button box. After excluding one participant for not following instructions and one outlier (see below), the final sample consisted of $N = 80$ participants (70 women; $M_{\text{age}} = 22.6$, $SD_{\text{age}} = 4.0$).

Participants were randomly assigned to the cells of a 2 (Type of primary task: depleting vs. non-depleting) $\times$ 4 (Type of secondary task: auditory SST, Stroop task, flanker task, unsolvable anagrams) between-subjects design. As a primary task, participants in the experimental conditions completed a visual SST, while participants in the control conditions completed a simple word categorization task. After completing the primary task, participants were assigned to one of four secondary tasks that varied in similarity.

Tasks

Lexical categorization task (non-depleting condition)

In a simple lexical categorization task, participants had to identify words as substantives or adjectives by pressing a blue (substantive) or red (adjective) button, respectively (Fig. 1). This task was designed to be cognitively easy and not to require self-control. Accordingly, participants were instructed to take their time and respond as accurately as possible. Participants completed 20 practice trials and four sequential test blocks of 100 randomized trials each, for an average duration of 13.4 minutes ($SD = 0.9$).

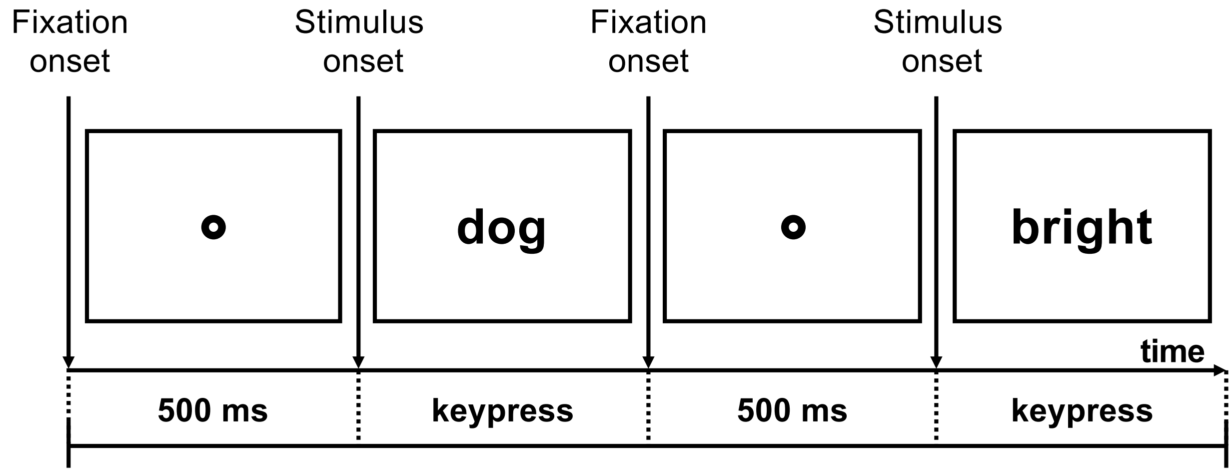

Figure 1

Time course of the simple categorization task. The substantives and adjectives remained on screen until individuals pressed a button. Each trial was preceded by a fixation point that was presented for 500 ms.

Visual SST (depleting condition)

We used a visual SST adapted from Enge et al. (2014), in which participants were instructed to discriminate between the words “left” and “right” by pressing a blue (“left”) or red (“right”) button (Fig. 2). However, they were also instructed to inhibit their response if the word changed its color from black to green (stop signal). In such signal trials, the stop signal occurred within the first 500 ms of the stimulus presentation. That is, the stop-signal delay (SSD) varied between 0 ms and a maximum of 500 ms. Importantly, it was adjusted on a trial-by-trial basis, depending on participants’ individual performance in order to adjust task difficulty according to the horse race model (Logan & Cowan, 1984). When participants correctly inhibited their response, the SSD increased by 50 ms in the next signal trial. In contrast, when participants failed to inhibit their response in a signal trial, the SSD decreased by 50 ms.

In addition, we used a performance-based feedback system inspired by Ridderinkhof et al. (1999) that was designed to discourage waiting strategies. In fact, some earlier studies have found progressive slowing in participants' response time as a strategy to avoid mistakes (Matzke et al., 2018). Importantly, such waiting strategies would also affect the extent of participants' response conflict and engagement because response inhibition becomes easier as participants proactively wait for stop signals. Thus, our main reason for the feedback system was to sustain and maximize participants' conflict between fast and accurate responding to increase task engagement (Dang & Hagger, 2019; cf. Wright et al., 2019). Specifically, after each trial, participants received a feedback in the form of point scores. The faster participants responded to the stimuli in a go trial relative to their mean response time, the higher the point score they received (ranging from 2 to 5 points). If the response time in a go trial exceeded the mean response time, participants received 2 points. However, when the response time in a go trial remained below the mean response time, points were awarded as follows: a ratio ranging from 0.9 to 1 between the response time in a single go trial and the mean response time were awarded 3 points (e.g., a response time in a single go trial of 660 ms and a mean response time of 700 ms equals a ratio of 0.94). A ratio between 0.8 and 0.9 led to 4 points and a ratio below 0.8 was awarded the maximum of 5 points. Following a false response in a go trial, participants received 1 point and 0 points if they did not respond at all. In signal trials, participants received 5 points upon correctly inhibiting their response and 1 point when they failed to inhibit their response.

Following a 20-trial practice block, participants completed four sequential test blocks of 100 randomized trials each (25% signal trials), for an average duration of 20.1 minutes ($SD = 0.2$). At the end of every block, participants received feedback including their mean response time on go trials and total point score across all go trials for the respective block.

Auditory SST (modality transfer)

The auditory SST (Fig. 3) was identical to the visual SST with the exception that a short sound (440 Hz, 100 ms) featured as stop-signal. Moreover, participants did not receive feedback on their performance. After a 20-trial practice block, participants

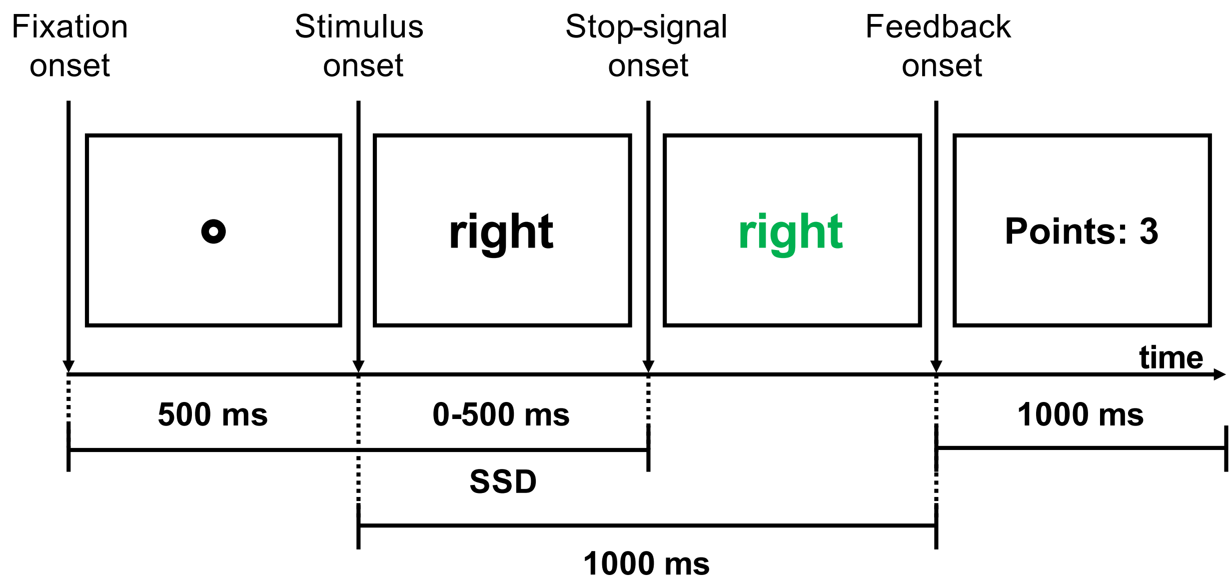

Figure 2

Time course of the visual stop-signal task for a go trial and a signal trial (25% rate). Target stimuli were presented after a 500 ms fixation point and displayed for a total of 1000 ms. In signal trials, a stop signal occurred between 0 and 500 ms. The initial stop-signal delay (SSD) was set to 250 ms and 50 ms were subtracted or added, respectively, according to the individual's performance. If participants mistakenly pressed a button during a signal trial, the SSD decreased by 50 ms. In contrast, the SSD was increased by 50 ms when the response was correctly inhibited. Note that the grey word (stop signal) was actually shown in green in the experiment.

completed four sequential test blocks of 100 randomized trials each (25% signal trials), for an average duration of 12.3 minutes ($SD = 0.2$). The key measure was the stop-signal response time (SSRT, Logan & Cowan, 1984), which expresses the speed of the inhibitory process after the onset of the stop signal. Higher SSRT values reflected lower inhibitory control and thus lower performance.

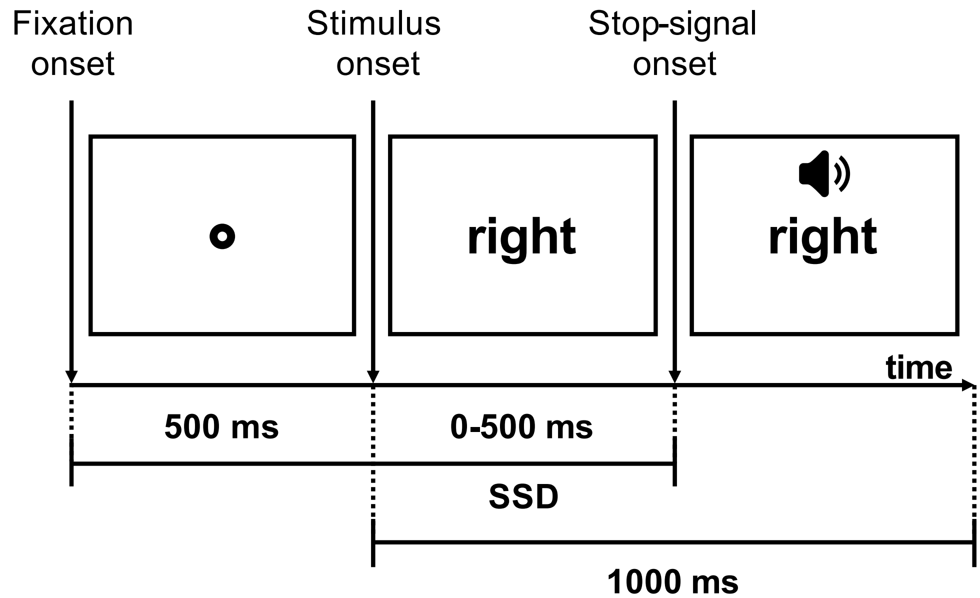

Figure 3

Time course of the auditory stop-signal task. The stimuli were displayed for 1000 ms and preceded by a fixation point that was presented for 500 ms. The initial stop-signal delay and its respective adjustment to the individual's performance was identical to the visual stop-signal task.

Stroop task (near transfer)

We conducted a color-naming Stroop task (Fig. 4) presenting three color names ("red," "blue," and "green") successively in varying font colors (red, blue, and green). Participants had to indicate the font color by pressing the appropriate color-coded button. Half of the trials were congruent (e.g., "red" written in red font color) and the other half were incongruent (e.g., "green" written in blue font color). Participants completed four sequential blocks of 96 randomized trials each, preceded by a 24-trial practice block. The task lasted for an average duration of 9.2 minutes ($SD = 1.0$). The key measure in this task was a Stroop interference score, which we calculated on the basis of the D scoring algorithm (Ebersole et al., 2016), where higher scores reflected lower inhibitory control and thus lower performance.

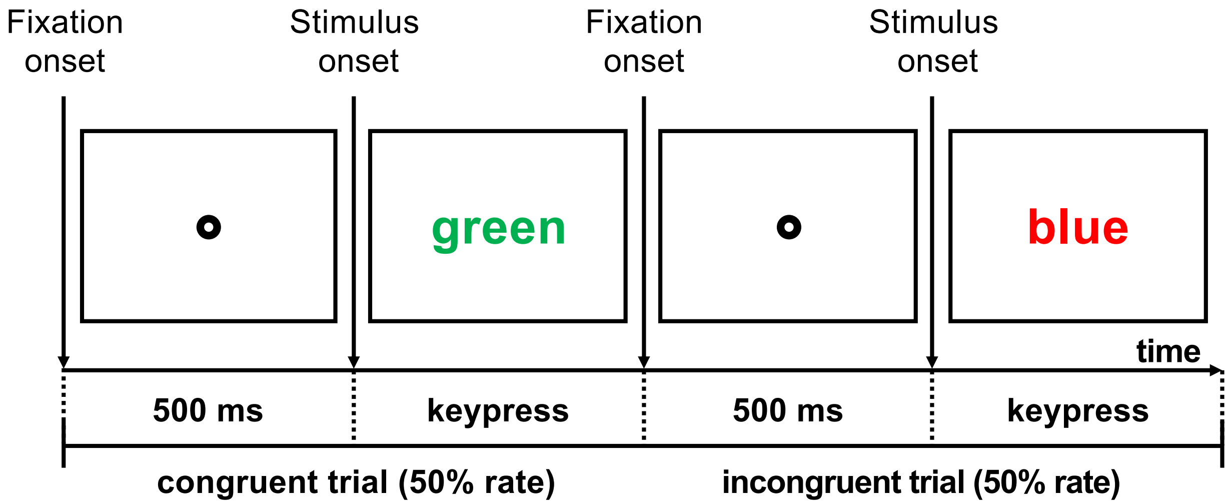

Figure 4

Time course of the Stroop task. The stimuli were pseudo-randomized ensuring that a trial's font color never coincided with the previous trial's font color or word meaning. At the beginning of every trial a fixation point was presented for 500 ms followed by the stimulus. Upon a keypress, the trial was completed and the next one began.

Eriksen flanker task (far transfer)

The Eriksen flanker task (Fig. 5) was adapted from Alderman et al. (2015) and consisted of a centrally positioned letter (target letter) and four flanking letters that either matched (congruent trial) or deviated (incongruent trial) from the target letter. Participants were asked to focus on the target letter while ignoring the other flanking letters and to press the blue button when "S" was the target letter and the red button when "H" was the target letter. After a 20-trial practice block, participants completed four sequential test blocks of 100 randomized trials each (25% incongruent trials), for an average duration of 22.4 minutes (SD = 0.2). Our dependent measure (flanker interference) reflected the difference between the μ parameters of the ex-Gaussian fitted response times of (correct) incongruent and congruent trials.

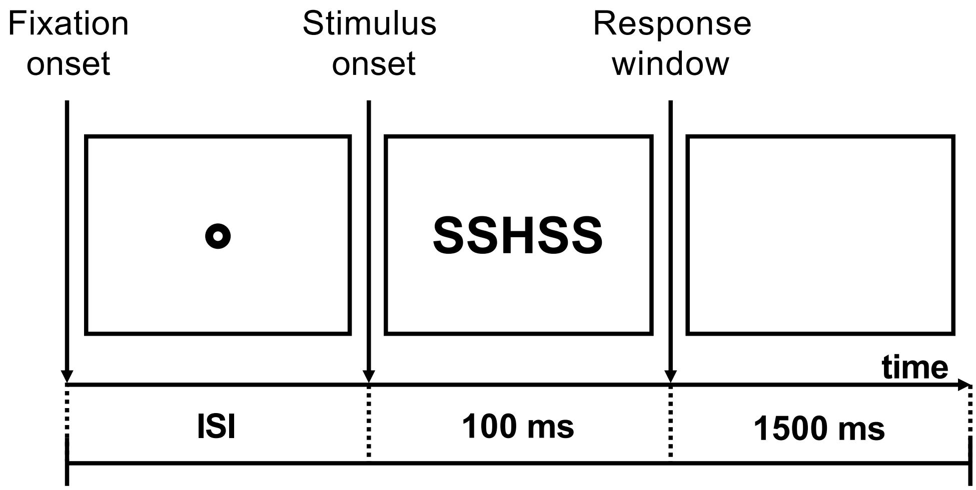

Figure 5

Time course of the Eriksen flanker task. Stimuli were displayed for 100 ms, followed by a response window of 1500 ms. In order to reduce expectancy effects/anticipatory responses, we used a random inter-stimulus interval (ISI) of 1100, 1300 or 1500 ms between the fixation point and the stimuli.

Unsolvable anagrams (meta-cognitive transfer)

In this task, participants were presented with a set of scrambled letters (six anagrams in total) and instructed to rearrange the letters in order to form meaningful words. Each anagram was displayed in capital letters in the upper center of the screen, and responses could be entered in an empty text box displayed below by using the keyboard. Participants were led to believe that all anagrams were solvable. In reality, only the first two anagrams were solvable, but not the other four. We asked participants to take their time to try and solve each anagram. We further told them to leave the box empty and move on to the next trial only if they were not able to solve the anagram even after thoughtful consideration. We used existing German anagrams from Topolinski et al. (2016) that increased in the number of letters. Solvable anagrams included KORDOT (DOKTOR) and GNUPFIM (IMPFUNG), whereas unsolvable anagrams included WAGRATU, RERUBEL, NOTLINFU, and FECHLIREN. Prior to the six test anagrams, participants completed two solvable practice anagrams. Persistence on the anagram task served as the key measure of self-control performance, calculated as the overall time spent on the four unsolvable test anagrams.

Manipulation check measures

After both the primary and the secondary task, participants completed single-item measures of difficulty ("How difficult did you find the task?"; 1 = *very easy*, 4 = *very difficult*), effort ("How much effort did you put into the task?"; 1 = *no effort*, 4 = *extreme effort*), fatigue ("How tired do you feel after doing the task?"; 1 = *not at all tired*, 4 = *extremely tired*), frustration ("Did you feel frustrated while you were doing the task?"; 1 = *not at all frustrated*, 4 = *very frustrated*), task enjoyment ("How much did you enjoy the task?"; 1 = *not at all*, 4 = *a lot*), and mood ("How do you feel right now?"; 1 = *very bad*, 4 = *very good*). In addition, participants reported whether they had experienced mixed feelings, conflict, and indecision during the task on a scale ranging from 1 (*not at all*) to 4 (*extremely*). These three ratings were averaged to form a measure of subjective ambivalence ($\alpha = .80$, Priester & Petty, 1996).

Data exclusions

Following our preregistration, we discarded all RTs below 200 ms and, in the case of the Stroop task, above 5000 ms. For the flanker task, we additionally removed RTs that were slower or faster than the participant's mean RT ± 3 standard deviations, respectively, to account for the sensitivity of the ex-Gaussian distribution to outliers (Spieler et al., 2000). No participant had to be excluded on account of overall error rates in any of the tasks (all error rates $< 22.3\%$). However, one participant was excluded because the standardized performance score (see Results) fell outside 1.5 IQR from the lower or upper quartile of the overall distribution (i.e., Tukey boxplot).

Results

We first conducted a manipulation check, assessing participants' responses on our additional measures after the first task. Summary statistics and results of all two-sided t-tests are presented in Table 2. Comparing the two conditions, we found no significant mean difference on mood or task enjoyment, suggesting that any effect would not have been emotionally driven. In addition, the conditions did not differ on reported fatigue, possibly because the somewhat boring nature of control conditions may also be fatiguing due to shifts in participants' motivational states (e.g., Brewer et al., 2017). However, we found significant differences on all other measures. That is, directly after the primary task, participants in the depleting condition reported higher difficulty, effort, frustration, and ambivalence. These findings were all supported by Bayesian re-analyses (Tab. 2). Thus, the results show that our manipulation was successful in that the visual SST was generally perceived as more arduous than the simple word categorization task.

Our hypothesis predicted a crossover interaction between type of primary (T1) and secondary task (T2) such that facilitation effects would turn into ego depletion effects as the similarity between T1 and T2 went from highly similar to highly dissimilar. We tested the predicted interaction pattern by means of a one-degree-of-freedom contrast test (Abelson & Prentice, 1997), for which the (T1, T2) contrast weights were specified as (2, -2) for the auditory SST, (1, -1) for the Stroop task, (-1, 1) for the flanker task, and (-2, 2) for the anagram task. We followed up this contrast with a test of the residual interaction (Abelson & Prentice, 1997). For the sake of comparability, all performance scores were z-standardized and, with the exception of persistence on the anagram task, reversed, such that higher scores would indicate better performance.

A 2 (primary) × 4 (secondary) ANOVA on the z-standardized performance measures revealed the predicted interaction (Table 3). The a priori contrast was significant, $p = .004, d = 0.61$; the residual interaction, however, was not, $p = .250$. Figure 6 shows the

Table 2

Summary Statistics, Two-Sided t Tests, and Two-Sided Default JSZ Bayes factors for All Manipulation Check Measures (Pilot Study)

	Condition	Mean	*SD*	*t*	*p*	BF_{10}
Mood	Depleting	3.05	0.50	-0.02	.981	2.30×10^{-1}
	Non-depleting	3.05	0.46			
Task enjoyment	Depleting	2.29	0.64	1.62	.109	7.30×10^{-1}
	Non-depleting	2.05	0.69			
Difficulty	Depleting	2.49	0.51	7.13	$< .001$	2.13×10^{7}
	Non-depleting	1.56	0.64			
Effort	Depleting	2.56	0.55	4.84	$< .001$	2.73×10^{3}
	Non-depleting	1.92	0.62			
Fatigue	Depleting	2.20	0.78	0.09	.926	2.30×10^{-1}
	Non-depleting	2.18	0.72			
Frustration	Depleting	2.07	0.61	5.13	$< .001$	6.69×10^{3}
	Non-depleting	1.41	0.55			
Ambivalence	Depleting	2.04	0.73	5.22	$< .001$	7.60×10^{3}
	Non-depleting	1.35	0.41			

Note. Bayes factors were calculated based on a JZS Cauchy prior for the effect size (with a scaling parameter of $r = \sqrt{2}/2$ in line with Rouder et al., 2012). $n_{\mathrm{non-depleting}} = 39$, $n_{\mathrm{depleting}} = 41$.

Table 3

Analysis of Variance with A Priori Contrast (Pilot Study)

Source	*SS*	*df*	*MS*	*F*	*p*
Task 1 (T1)	0.19	1	0.19	0.22	.639
Task 2 (T2)	0.31	3	0.10	0.12	.945
T1 x T2	8.46	3	2.82	3.38	.023
A priori contrast	6.10	1	6.10	7.56	.004
Residual interaction	2.36	2	1.18	1.42	.250
Error	60.02	72	0.83		

Note. SS = Sum of squares; MS = Mean square.

obtained interaction pattern: High similarity between the primary and secondary task led to a facilitation effect, whereas low similarity resulted in an ego depletion effect. In Bayesian terms, if we consider the t-value (square root of the F-value of our a priori contrast) with cell sizes of 41 and 39 participants, we obtain a high and robust Bayes Factor of 11.37 (with a default JZS Cauchy prior width of $r = \sqrt{2}/2$) indicating strong evidence for our hypothesis (cf. Jeffreys, 1961).

Our main conclusion does not depend on the exclusion of outliers. That is, after including one outlier identified in a boxplot, the a priori contrast was still significant, $p = .010, d = 0.53$ (although the removal of the outlier decreased the p-value of the residual interaction so that it would be considered significant according to our preregistered alpha criterion). Moreover, including the duration of the secondary task as a covariate did not change our results, ruling out the possibility that our effect may have been confounded with task duration.

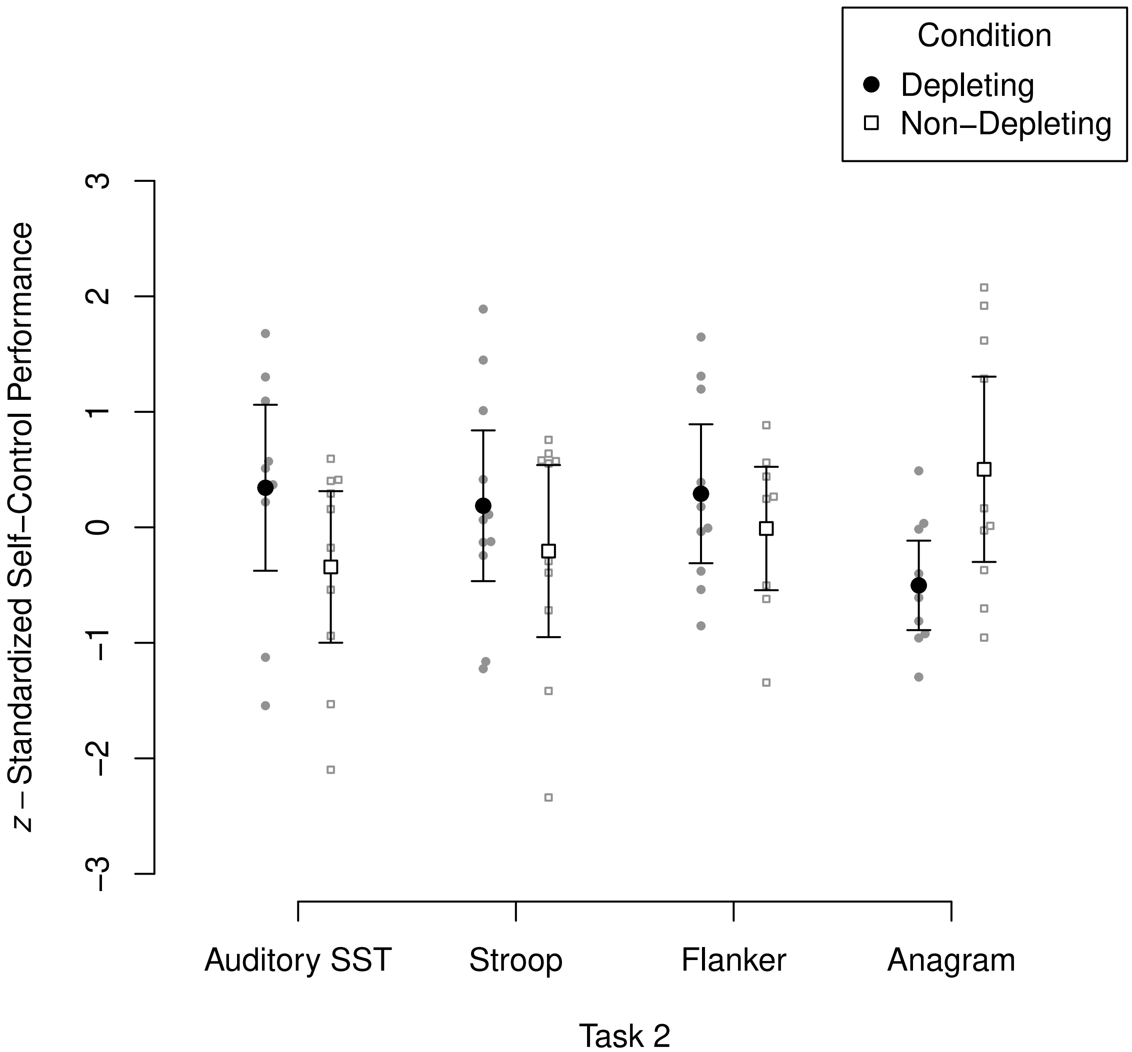

Figure 6

Beeplot of z-standardized performance scores in the eight groups in the pilot study. Positive values indicate better performance. Small points represent individual observations, large points represent group means, and error bars represent 95% confidence intervals. SST = stop-signal task.

Discussion

We obtained preliminary evidence for the moderating role of task similarity on the ego depletion effect. Conceptually replicating the findings of Dewitte et al. (2009), our results suggest that ego depletion effects may be more likely to occur between dissimilar rather than similar tasks. Going beyond Dewitte et al. (2009), we used experimental tasks more typical of the ego depletion literature. Furthermore, our manipulation of task similarity was based on a fine-grained taxonomy of the executive functions required by self-control tasks, allowing for a systematic ordering of tasks from most to least similar. The effect sizes tracked this ordering very well, yielding the predicted crossover interaction with facilitation effects turning into ego depletion effects as task similarity decreased. Our close adherence to the preregistered analysis plan should bolster our confidence that our finding of a crossover interaction is not a false positive. Nevertheless, any single study obviously cannot do much more than invite replication. Given the encouraging first results and the high theoretical and practical value of confirming or disconfirming the results with new data, we proposed the following registered study in an attempt to shed more light on the true value of our findings.

Registered study

This registered study was a direct replication of the pilot study. All materials, data, and analysis scripts are available through the same repository (https://osf.io/b9yek/). The goal was to conduct a high-powered test of the hypothesis that task similarity moderates the ego depletion effect in a sequential-task paradigm. In other words, we expected an ego depletion effect to occur between dissimilar tasks (i.e., the visual SST and the anagram task) and a facilitation effect between similar tasks (i.e., the visual SST and the auditory SST). In light of the possibility that a positive replication result could emerge due to unknown characteristics of the original study site, which would have remained constant, we recruited participants at two research sites using different experimenters and participant pools.

Modifications to the pilot study methods

We adhered to the methods presented above with the exception of two important changes. First, in order to increase comparability between the dependent tasks, we extended the ex-Gaussian approach used to derive the dependent measure for the flanker task to the Stroop task. Thus, instead of the D score, we calculated the difference between the μ parameters of the ex-Gaussian fitted response times of (correct) incongruent and congruent trials as our main dependent measure for the Stroop task. Applying this change to the analysis of our pilot data yielded virtually the same results. Second, we sought to make the duration of the different tasks more similar by shortening the flanker task from four to two blocks of 100 trials each.

Planned sample size

We conducted a safeguard power analysis (Perugini et al., 2014) to determine the sample size that affords at least 90% power for detecting the predicted crossover interaction as defined by the a priori contrast. Safeguard power analysis is a conservative procedure which does not use the estimate of the population effect size from the pilot study, but the lower limit of a confidence interval for the contrast effect obtained in the main analysis of the pilot study (after the exclusion of outliers). We used an 80% confidence interval, which had a lower limit of $d_{\mathrm{c_{ll}}} = 0.42$. In other words, our power analysis rested upon the assumption that the effect we aimed to detect is not smaller than $d_{\mathrm{c_{ll}}} = 0.42$, an assumption that could be made with a high degree of confidence. This analysis indicated that a total of $N = 298$ participants would be required for a 90%-powered one-sided significance test with an alpha of 1% for the proposed study. Anticipating that some participants might fail to meet our inclusion criteria, we finally decided to collect a total sample size of $N = 310$.

Participants and experimental setting

Between the months of October, 2020 and March, 2021, we recruited a total of $N = 314$ students from a Swiss (CH; $n = 158$) and a German (DE; $n = 156$) university in exchange for course credit or a small monetary reward (CHF 15 in CH, EUR 10 in DE). Participants were tested in group sessions with up to four participants. All tasks and control measures were presented on a laptop (CH) or desktop computer (DE) running the experiment software OpenSesame (Mathôt et al., 2012).[1] All responses and response times (RT) were logged and measured with millisecond precision using a labeled button box (CH) or a keyboard (DE). Following our preregistration, we had to exclude a total of 14 participants who either experienced technical errors ($n = 7$), did not meet eligibility criteria (i.e., age > 35; $n = 1$), exceeded the accepted overall error rate of 40% on any of the secondary inhibitory control tasks ($n = 1$), or were identified as outliers with respect to the primary dependent variable ($n = 5$), resulting in a final sample of $N = 300$ participants (219 women; $M_{\text{age}} = 22.7$, $SD_{\text{age}} = 4.0$).

Analysis plan

With the exceptions noted above, we followed the same analysis steps and applied the same exclusion criteria as reported for the pilot study. One issue that was raised in the review process was the use of a contrast analysis. A possible concern is that this test makes the strong assumption of linearity. There were indeed no strong theoretical grounds on which to assume that the effect sizes should change in linear fashion across levels of similarity. However, it was the simplest pattern to assume and the validity of our focused comparison does not depend on this assumption. If effect sizes would have changed non-linearly, we would have detected this visually as well as through the residual

[1] The experiment code was rewritten to conform with the then-current version of OpenSesame (3.3.3 *Lentiform Loewenfeld*, Mathôt et al., 2012), with which the registered experiment was run.

interaction. Given the greater statistical power of a single-degree-of-freedom test (compared with an omnibus test), we retained this type of analysis for the registered study.

In order to corroborate our results, however, we extended our analyses to different outcome measures. Specifically, since we changed our primary dependent measure for the Stroop task to align with our measure for the flanker task, we report our main analysis for both the modified and the original set of dependent measures. A detailed description of the adapted analysis plan can be found in our preregistration at https://osf.io/hzwk4.

Results

Manipulation check

Descriptive and inferential statistics for all manipulation check measures are presented in Table 4. There were significant differences between experimental conditions on all measures but fatigue. More specifically, compared to participants in the non-depleting condition, participants in the depleting condition reported higher difficulty, effort, frustration, and ambivalence after the first task. Interestingly, participants in the depleting condition reported worse mood but higher task enjoyment compared to participants in the non-depleting condition. Again, these findings were all supported by corresponding Bayesian analyses (Tab. 4). In sum, the results suggest that the manipulation was successful in that the visual SST was generally perceived as more arduous than the simple word categorization task.

Confirmatory analysis

A 2 (type of primary task) $\times$ 4 (type of secondary task) ANOVA on the z-standardized performance measures revealed the predicted interaction (Table 5). The a priori contrast was significant, $p < .001, d = 0.43$, whereas the residual interaction was not, $p = .907$. Figure 7 shows the obtained interaction pattern, suggesting that high similarity

Table 4

Summary Statistics, Two-Sided t Tests, and Two-Sided Default JSZ Bayes factors for All Manipulation Check Measures (Registered Study)

	Condition	Mean	*SD*	*t*	*p*	BF_{10}
Mood	Depleting	2.93	0.50	-3.11	.002	1.26×10^{1}
	Non-depleting	3.12	0.55			
Task enjoyment	Depleting	2.37	0.76	6.65	$< .001$	4.87×10^{7}
	Non-depleting	1.83	0.65			
Difficulty	Depleting	2.55	0.66	20.40	$< .001$	7.47×10^{53}
	Non-depleting	1.24	0.45			
Effort	Depleting	2.47	0.55	11.48	$< .001$	1.46×10^{22}
	Non-depleting	1.77	0.51			
Fatigue	Depleting	2.08	0.66	1.69	.093	4.90×10^{-1}
	Non-depleting	1.96	0.62			
Frustration	Depleting	2.14	0.59	10.18	$< .001$	8.02×10^{17}
	Non-depleting	1.44	0.59			
Ambivalence	Depleting	2.12	0.58	11.86	$< .001$	1.70×10^{23}
	Non-depleting	1.42	0.45			

Note. Bayes factors were calculated based on a JZS Cauchy prior for the effect size (with a scaling parameter of $r = \sqrt{2}/2$ in line with Rouder et al., 2012). $n_{\text{non-depleting}} = 145$, $n_{\text{depleting}} = 155$.

Table 5

Analysis of Variance with A Priori Contrast (Registered Study)

Source	SS	*df*	MS	F	p
Task 1 (T1)	7.29	1	7.29	9.50	.002
Task 2 (T2)	0.68	3	0.23	0.30	.829
T1 x T2	11.17	3	3.72	4.85	.003
A priori contrast	11.02	1	11.02	14.12	$< .001$
Residual interaction	0.15	2	0.07	0.10	.907
Error	224.11	292	0.77		

Note. SS = Sum of squares; MS = Mean square.

between the primary and secondary task led to facilitation effects, whereas the lowest level of similarity reflected a reversed, potential ego depletion effect. Moreover, a Bayesian t test of our a priori contrast and cell sizes of 155 and 145 participants revealed a Bayes factor of 191.63 (based on a one-sided JZS Cauchy prior with a scaling parameter of $r = \sqrt{2}/2$), indicating decisive evidence in favor of our hypothesis (cf. Jeffreys, 1961).

Exploratory analyses

We conducted a number of robustness checks to further corroborate our findings (see supplementary analyses at https://osf.io/y37xv/, for more detailed results). First, the exclusion of outliers did not affect our results. In fact, after including the five outliers, the effect size of the a priori contrast increased, $p < .001$, $d = 0.45$, whereas the residual interaction remained non-significant, $p = .792$. Second, including duration of the secondary task as a covariate did not change our results, ruling out the possibility that our effect may have been confounded with task duration. Third, using the D score for the Stroop task as

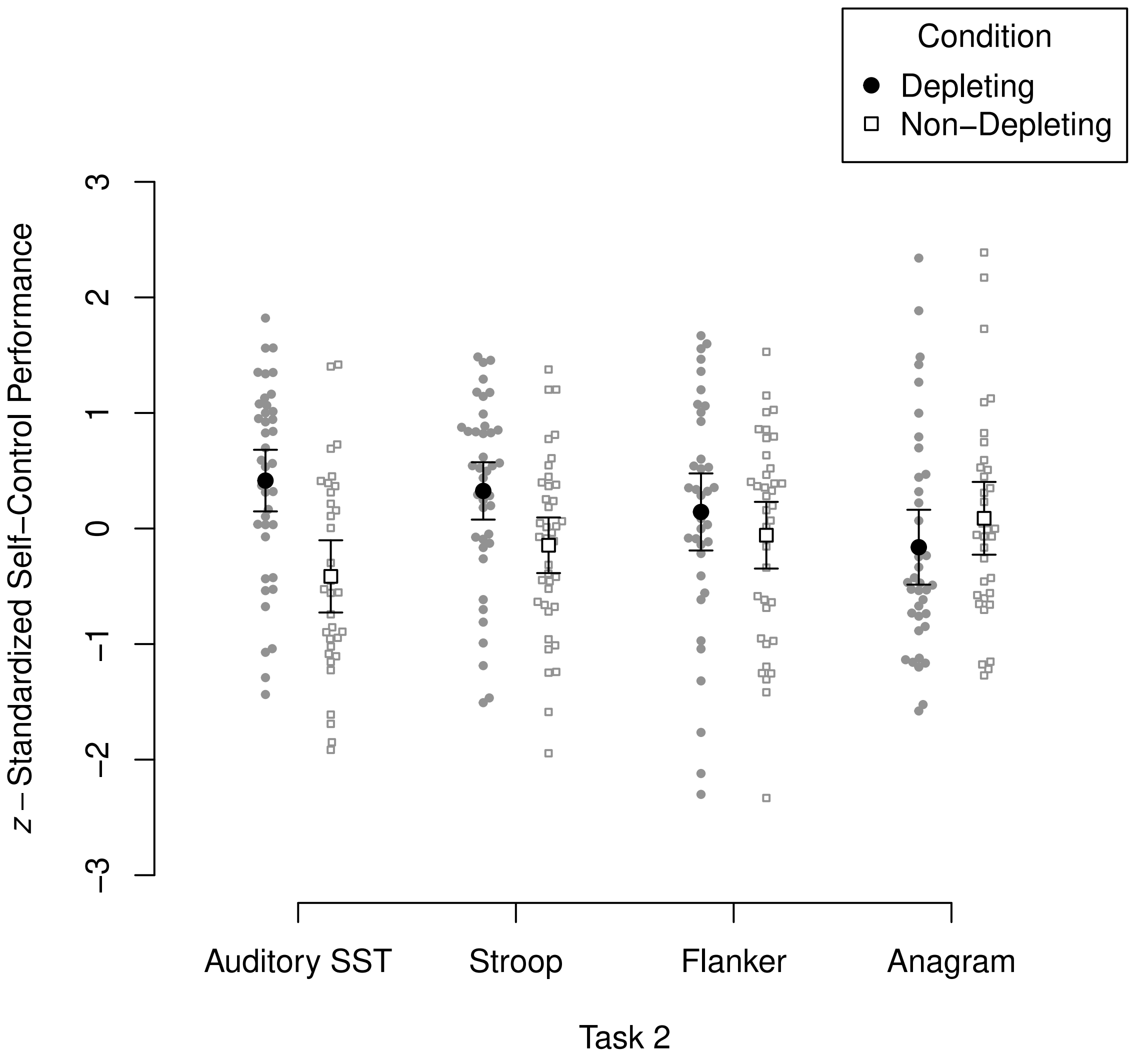

Figure 7

Beeplot of z-standardized performance scores in the eight groups in the registered study. Positive values indicate better performance. Small points represent individual observations, large points represent group means, and error bars represent 95% confidence intervals. SST = stop-signal task.

in the pilot study revealed virtually the same pattern of results. Fourth, we noticed that some participants had high error rates in the word categorization task and/or adopted all-or-nothing response strategies in the two SSTs (i.e., responding to all or none of the stop signals), indicating potential problems regarding their involvement or compliance with task instructions. As an exploratory analysis, we excluded participants who (1) exhibited overall error rates above 10% in the word categorization task ($n = 4$) or (2) responded to each or none of the stop signals in at least two experimental blocks of the visual ($n = 10$) or auditory SST ($n = 0$). This did not materially affect our results.

Finally, post-hoc pairwise comparisons indicated that, on the auditory SST, standardized performance was significantly higher for participants in the depleting condition ($M = 0.41$, $SD = 0.83$) than for participants in the non-depleting condition ($M = -0.41$, $SD = 0.88$), $t(292) = 4.02$, $p < .001$, confirming the presence of a facilitation effect. The same was true for the Stroop task, with standardized performance scores being significantly higher for participants in the depleting condition ($M = 0.33$, $SD = 0.79$) as compared to participants in the non-depleting condition ($M = -0.15$, $SD = 0.75$), $t(292) = 2.42$, $p = .016$. Comparisons were non-significant for the flanker task, $p = .320$, and the unsolvable anagram task, $p = .230$. In addition, corresponding Bayesian two-sided t tests (with a Cauchy prior width of $r = \sqrt{(2)}/2$) indicated decisive evidence for the hypothesis that participants in the depleting condition performed better on the auditory SST than participants in the non-depleting condition, $\text{BF}_{10} = 221.12$, as well as substantial evidence for the hypothesis that participants in the depleting condition performed better on the Stroop task than participants in the non-depleting condition, $\text{BF}_{10} = 5.71$. In contrast, for both the flanker task and unsolvable anagrams, the Bayes factors, $\text{BF}_{10} = 0.35$ and $\text{BF}_{10} = 0.42$, respectively, indicated only weak evidence in favor of the null hypothesis that posited no mean difference between the two conditions.

Discussion

The results of this registered study confirmed the findings from the pilot study in providing more compelling evidence of a moderating role of task similarity for the ego depletion effect. This was again reflected in a crossover interaction that emerged between primary and secondary tasks, with facilitation effects tending to reverse into an ego depletion effect as task similarity decreased. With regard to our taxonomy, follow-up tests further corroborated the presence of facilitation effects both between highly similar modality-transfer tasks (differing in the kind of stimuli) and—to a lesser extent—between moderately similar near-transfer tasks (differing in the kind of task domain/function). However, we did no find conclusive evidence of an ego depletion effect, even in the context of highly dissimilar tasks that required a meta-cognitive transfer from one task context to another.

General discussion

Much research has been devoted to understanding how self-control performance depends on the prior execution of self-control. While the ego depletion hypothesis predicts a negative effect, a number of studies have brought forth arguments and preliminary evidence for a positive, facilitating effect, suggesting that under certain circumstances self-control performance may improve (rather than deteriorate) from one task to another (e.g., Dang et al., 2014; Dewitte et al., 2009; Xiao et al., 2014).

The present research suggests that whether the effect is positive or negative may depend on the cognitive processes involved in consecutive self-control tasks. Based on theoretical accounts of executive functions involved in self-control, we presented a taxonomy of inhibitory self-control tasks that infers task similarity from the level of transferability of executive functions between the tasks. Accordingly, we hypothesized that task similarity should moderate the ego depletion effect, such that facilitation effects would be more likely to emerge between similar tasks, whereas ego depletion effects would be more likely to emerge between dissimilar tasks.

Replicating results from a pilot study, a high-powered registered study provides strong and robust evidence for the hypothesized moderation effect of task-similarity. Specifically, our results revealed pronounced facilitation effects for similar tasks, such that participants who engaged in two highly similar cognitive inhibition tasks showed significantly better performance in the second task compared to participants who first engaged in a cognitively undemanding and unrelated task. Importantly, these facilitation effects declined with decreasing task similarity, reversing in direction at the highest level of dissimilarity (though not significantly so).

The results lend support to the idea that the paradigm used to study ego depletion can actually produce facilitation effects, as opposed to ego depletion effects, under conditions that allow people to adjust to task demands (Dang et al., 2013; Dewitte et al., 2009). Our research goes beyond previous findings by elaborating on the role of task similarity as a determining factor in the context of self-regulatory after-effects. In other words, whether adapting to specific task demands in a previous task ultimately enhances people's self-control performance in a subsequent task should depend on the extent to which the two tasks overlap in terms of the underlying psychological processes.

Attending to the cognitive processes underlying self-regulatory behavior allows for the development of fine-grained task taxonomies, along which self-control tasks can be categorized, compared, and selected. Our present work provides compelling empirical support for a first attempt at organizing commonly used cognitive inhibition tasks and illustrates how this taxonomy may be used to ascertain task similarity. Adapting and extending this taxonomy to different underlying processes and associated tasks could considerably advance future research by enabling more principled task selection and effective experimental designs.

Although the size of facilitation effects decreased with decreasing task similarity, we did not obtain significant ego depletion effects for dissimilar tasks. One explanation may be drawn from recently accumulating evidence that casts doubt on the existence (Evan C. Carter et al., 2015; Evan C. Carter & McCullough, 2014; Etherton et al., 2018; Hagger et al., 2016; Vohs et al., 2021) or, at least, the size of the effect (Dang et al., 2021; Dang, 2018; Francis et al., 2018; Garrison et al., 2019; Lin et al., 2020). According to this interpretation, our results may either provide additional evidential support against the existence of ego depletion or, assuming that the effect is real but small, our design might have lacked sufficient power to detect it within a single experimental condition. If we assume that the processes leading to facilitation and depletion effects are likely to co-occur, with their relative strength depending on task similarity, then it makes sense to expect depletion effects that are close to zero when tasks are relatively similar. This raises the possibility that there might be even more dissimilar tasks than SST and unsolvable anagrams, which would produce more reliable ego depletion effects.

Our research strongly suggests that ego depletion effects are not unitary and will vary from task to task. Perhaps the lack of support for the ego depletion hypothesis is due to the specific tasks we used (or their unique combination). Indeed, a recent reanalysis of the large Many Labs 3 data set (Vadillo et al., 2018) found no indication of an ego depletion effect in terms of persistence on an (impossible) anagram task after engaging in a Stroop task. In comparison, it is worth noting, that the considered paradigm involved a much shorter Stroop task of 63 trials (vs. 400 SST trials in the present study) as depleting task and only one, time-limited anagram (vs. four self-paced anagrams in the present study) as dependent task. More importantly, though, a meta-analytic reassessment (including several comparable studies) by the same authors raised further doubt on the ability of unsolvable anagrams to measure depletion effects. Future work with dissimilar self-control tasks should therefore consider alternatives to unsolvable anagrams that are more reliable in capturing self-control performance.

To conclude, the present research suggests that it may be fruitful to further entertain the idea that self-control performance may not only decrease but also increase as a result of prior exercise of self-control. We proposed a simple theoretical framework for the development of task taxonomies that takes into account process-level executive functions involved in self-control. It is our hope that a better understanding of task similarity may allow for better experimental designs and more rigorous tests of self-control performance in consecutive tasks.

References

Abelson, R. P., & Prentice, D. A. (1997). Contrast Tests of Interaction Hypotheses. *Psychological Methods*, *2*(4), 315–328. https://doi.org/10.1037/1082-989X.2.4.315

Alderman, B. L., Olson, R. L., Bates, M. E., Selby, E. A., Buckman, J. F., Brush, C. J., Panza, E. A., Kranzler, A., Eddie, D., & Shors, T. J. (2015). Rumination in major depressive disorder is associated with impaired neural activation during conflict monitoring. *Frontiers in Human Neuroscience*, *9*(MAY), 269. https://doi.org/10.3389/fnhum.2015.00269

Allom, V., & Mullan, B. (2014). Individual differences in executive function predict distinct eating behaviours. *Appetite*, *80*, 123–130. https://doi.org/10.1016/j.appet.2014.05.007

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*(5), 1252–1265. https://doi.org/10.1037//0022-3514.74.5.1252

Boksem, M. A. S., Meijman, T. F., & Lorist, M. M. (2006). Mental fatigue, motivation and action monitoring. *Biological Psychology*, *72*(2), 123–132. https://doi.org/10.1016/j.biopsycho.2005.08.007

Boksem, M. A. S., & Tops, M. (2008). Mental fatigue: Costs and benefits. *Brain Research Reviews*, *59*(1), 125–139. https://doi.org/10.1016/j.brainresrev.2008.07.001

Botvinick, M. M., Carter, C. S., Braver, T. S., Barch, D. M., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652. https://doi.org/10.1037/0033-295X.108.3.624

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences*, *8*(12), 539–546. https://doi.org/10.1016/j.tics.2004.10.003

Braem, S., Abrahamse, E. L., Duthoo, W., & Notebaert, W. (2014). What determines the specificity of conflict adaptation? A review, critical analysis, and proposed synthesis. *Frontiers in Psychology*, *5*(SEP). https://doi.org/10.3389/fpsyg.2014.01134

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*, *16*(2), 106–113. https://doi.org/10.1016/j.tics.2011.12.010

Brewer, G. A., Lau, K. K. H., Wingert, K. M., Hunter Ball, B., & Blais, C. (2017). Examining depletion theories under conditions of within-task transfer. *Journal of Experimental Psychology: General*, *146*(7), 988–1008. https://doi.org/10.1037/xge0000290

Carter, Evan C., Kofler, L. M., Forster, D. E., & McCullough, M. E. (2015). A series of meta-analytic tests of the depletion effect: Self-control does not seem to rely on a limited resource. *Journal of Experimental Psychology: General*, *144*(4), 796–815. https://doi.org/10.1037/xge0000083

Carter, Evan C., & McCullough, M. E. (2013). Is ego depletion too incredible? Evidence for the overestimation of the depletion effect. *Behavioral and Brain Sciences*, *36*(6), 683–684. https://doi.org/10.1017/S0140525X13000952

Carter, Evan C., & McCullough, M. E. (2014). Publication bias and the limited strength model of self-control: has the evidence for ego depletion been overestimated? *Frontiers in Psychology*, *5*, 1–11. https://doi.org/10.3389/fpsyg.2014.00823

Chatzisarantis, N. L. D., & Hagger, M. S. (2015). Illusionary delusions. Willingness to exercise self-control can mask effects of glucose on self-control performance in experimental paradigms that use identical self-control tasks. *Appetite*, *84*, 322–324. https://doi.org/10.1016/j.appet.2014.10.025

Dang, J. (2018). An updated meta-analysis of the ego depletion effect. *Psychological Research*, *82*(4), 645–651. https://doi.org/10.1007/s00426-017-0862-x

Dang, J., Barker, P., Baumert, A., Bentvelzen, M., Berkman, E., Buchholz, N., Buczny, J., Chen, Z., De Cristofaro, V., Vries, L. de, Dewitte, S., Giacomantonio, M., Gong, R., Homan, M., Imhoff, R., Ismail, I., Jia, L., Kubiak, T., Lange, F., … Zinkernagel, A. (2021). A Multilab Replication of the Ego Depletion Effect. *Social Psychological and Personality Science*, *12*(1), 14–24. https://doi.org/10.1177/1948550619887702

Dang, J., Dewitte, S., Mao, L., Xiao, S., & Shi, Y. (2013). Adapting to an initial self-regulatory task cancels the ego depletion effect. *Consciousness and Cognition*, *22*(3), 816–821. https://doi.org/10.1016/j.concog.2013.05.005

Dang, J., & Hagger, M. S. (2019). Time to Set a New Research Agenda for Ego Depletion and Self-Control. *Social Psychology*, *50*(5-6), 277–281. https://doi.org/10.1027/1864-9335/a000399

Dang, J., Xiao, S., & Dewitte, S. (2014). Self-control depletion is more than motivational switch from work to fun: The indispensable role of cognitive adaptation. *Frontiers in Psychology*, *5*(AUG), 933. https://doi.org/10.3389/fpsyg.2014.00933

Dewitte, S., Bruyneel, S., & Geyskens, K. (2009). Self-Regulating enhances self-regulation in subsequent consumer decisions involving similar response conflicts. *Journal of Consumer Research*, *36*(3), 394–405. https://doi.org/10.1086/598615

Ebersole, C. R., Atherton, O. E., Belanger, A. L., Skulborstad, H. M., Allen, J. M., Banks, J. B., Baranski, E., Bernstein, M. J., Bonfiglio, D. B. V., Boucher, L., Brown, E. R., Budiman, N. I., Cairo, A. H., Capaldi, C. A., Chartier, C. R., Chung, J. M., Cicero, D. C., Coleman, J. A., Conway, J. G., … Nosek, B. A. (2016). Many Labs 3: Evaluating participant pool quality across the academic semester via replication. *Journal of Experimental Social Psychology*, *67*, 68–82. https://doi.org/10.1016/j.jesp.2015.10.012

Egner, T. (2014). Creatures of habit (and control): A multi-level learning perspective on the modulation of congruency effects. *Frontiers in Psychology*, *5*(NOV), 1247. https://doi.org/10.3389/fpsyg.2014.01247

Eisenberg, I. W., Bissett, P. G., Zeynep Enkavi, A., Li, J., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Uncovering the structure of self-regulation through data-driven ontology discovery. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-10301-1

Enge, S., Behnke, A., Fleischhauer, M., Küttler, L., Kliegel, M., & Strobel, A. (2014). No evidence for true training and transfer effects after inhibitory control training in young healthy adults. *Journal of Experimental Psychology: Learning Memory and Cognition*, *40*(4), 987–1001. https://doi.org/10.1037/a0036165

Enriquez-Geppert, S., Huster, R. J., & Herrmann, C. S. (2013). Boosting brain functions: Improving executive functions with behavioral training, neurostimulation, and neurofeedback. *International Journal of Psychophysiology*, *88*(1), 1–16. https://doi.org/10.1016/j.ijpsycho.2013.02.001

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. https://doi.org/10.3758/BF03203267

Etherton, J. L., Osborne, R., Stephenson, K., Grace, M., Jones, C., & De Nadai, A. S. (2018). Bayesian analysis of multimethod ego-depletion studies favours the null hypothesis. *British Journal of Social Psychology*, *57*(2), 367–385. https://doi.org/10.1111/bjso.12236

Francis, Z., Milyavskaya, M., Lin, H., & Inzlicht, M. (2018). Development of a Within-Subject, Repeated-Measures Ego-Depletion Paradigm: Inconsistent Results and Future Recommendations. *Social Psychology*, *49*(5), 271–286. https://doi.org/10.1027/1864-9335/a000348

Friedman, N. P., & Miyake, A. (2004). The Relations Among Inhibition and Interference Control Functions: A Latent-Variable Analysis. *Journal of Experimental Psychology: General*, *133*(1), 101–135. https://doi.org/10.1037/0096-3445.133.1.101

Friese, M., Loschelder, D. D., Gieseler, K., Frankenbach, J., & Inzlicht, M. (2019). Is Ego Depletion Real? An Analysis of Arguments. *Personality and Social Psychology Review*, *23*(2), 107–131. https://doi.org/10.1177/1088868318762183

Garrison, K. E., Finley, A. J., & Schmeichel, B. J. (2019). Ego Depletion Reduces Attention Control: Evidence From Two High-Powered Preregistered Experiments. *Personality and Social Psychology Bulletin*, *45*(5), 728–739. https://doi.org/10.1177/0146167218796473

Gärtner, A., & Strobel, A. (2021). Individual Differences in Inhibitory Control: A latent Variable Analysis. *Journal of Cognition*, *4*(1), 1–18. https://doi.org/10.5334/joc.150

Giboin, L. S., & Wolff, W. (2019). The effect of ego depletion or mental fatigue on subsequent physical endurance performance: A meta-analysis. *Performance Enhancement and Health*, *7*(1-2), 23. https://doi.org/10.1016/j.peh.2019.100150

Gonthier, C., Braver, T. S., & Bugg, J. M. (2016). Dissociating proactive and reactive control in the Stroop task. *Memory and Cognition*, *44*(5), 778–788. https://doi.org/10.3758/s13421-016-0591-1

Hagger, M. S., Chatzisarantis, N. L. D., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., Brand, R., Brandt, M. J., Brewer, G. A., Bruyneel, S., Calvillo, D. P., Campbell, W. K., Cannon, P. R., Carlucci, M., Carruth, N. P., Cheung, T., Crowell, A., De Ridder, D. T. D., Dewitte, S., … Zwienenberg, M. (2016). A multilab preregistered replication of the ego-depletion effect [Article]. *Perspectives on Psychological Science*, *11*(4), 546–573. https://doi.org/10.1177/1745691616652873

Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. D. (2010). Ego depletion and the strength model of self-control: A meta-analysis. *Psychological Bulletin*, *136*(4), 495–525. https://doi.org/10.1037/a0019486

Heller, S., Borsay, F., & Ullrich, J. (2017). Social power and dimensions of self-control: Does power benefit initiatory self-control but impair inhibitory self-control? *Cogent*

Psychology, *4*(1). https://doi.org/10.1080/23311908.2017.1288351

Herlambang, M. B., Taatgen, N. A., & Cnossen, F. (2019). The Role of Motivation as a Factor in Mental Fatigue. *Human Factors*, *61*(7), 1171–1185. https://doi.org/10.1177/0018720819828569

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford University Press.

Kato, Y., Endo, H., & Kizuka, T. (2009). Mental fatigue and impaired response processes: Event-related brain potentials in a Go/NoGo task. *International Journal of Psychophysiology*, *72*(2), 204–211. https://doi.org/10.1016/j.ijpsycho.2008.12.008

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching-a review. *Psychological Bulletin*, *136*(5), 849–874. https://doi.org/10.1037/a0019842

Lange, F. (2015). If ego depletion cannot be studied using identical tasks, it is not ego depletion. *Appetite*, *84*, 325–327. https://doi.org/10.1016/j.appet.2014.10.026

Lin, H., Saunders, B., Friese, M., Evans, N. J., & Inzlicht, M. (2020). Strong Effort Manipulations Reduce Response Caution: A Preregistered Reinvention of the Ego-Depletion Paradigm. *Psychological Science*, *31*(5), 531–547. https://doi.org/10.1177/0956797620904990

Lindsay, D. S., & Jacoby, L. L. (1994). Stroop Process Dissociations: The Relationship Between Facilitation and Interference. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 219–234. https://doi.org/10.1037/0096-1523.20.2.219

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: General and Special theory of an act of control. *Psychological Review*, *91*(3), 295–327. https://doi.org/10.1037/a0035230

Logan, G. D., & Zbrodoff, N. J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & Cognition*, *7*(3), 166–174. https://doi.org/10.3758/BF03197535

Lorist, M. M., Boksem, M. A. S., & Ridderinkhof, K. R. (2005). Impaired cognitive control and reduced cingulate activity during mental fatigue. *Cognitive Brain Research*, *24*(2), 199–205. https://doi.org/10.1016/j.cogbrainres.2005.01.018

Lurquin, J. H., & Miyake, A. (2017). Challenges to ego-depletion research go beyond the replication crisis: A need for tackling the conceptual crisis. *Frontiers in Psychology*, *8*(APR), 1–5. https://doi.org/10.3389/fpsyg.2017.00568

Maloney, P. W., Grawitch, M. J., & Barber, L. K. (2012). The multi-factor structure of the Brief Self-Control Scale: Discriminant validity of restraint and impulsivity. *Journal of Research in Personality*, *46*(1), 111–115. https://doi.org/10.1016/j.jrp.2011.10.001

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. https://doi.org/10.3758/s13428-011-0168-7

Matzke, D., Verbruggen, F., & Logan, G. D. (2018). The Stop-Signal Paradigm. In *Stevens' handbook of experimental psychology and cognitive neuroscience* (pp. 1–45). John Wiley & Sons, Inc. https://doi.org/10.1002/9781119170174.epcn510

Osgood, J. M. (2017). Effect of ego-depletion typing task on Stroop does not extend to diverse online sample. *Journal of Articles in Support of the Null Hypothesis*, *13*(2), 83–90.

Pattyn, N., Van Cutsem, J., Dessy, E., & Mairesse, O. (2018). Bridging exercise science, cognitive psychology, and medical practice: Is "cognitive fatigue" a remake of "the emperor's new clothes"? *Frontiers in Psychology*, *9*(SEP). https://doi.org/10.3389/fpsyg.2018.01246

Perugini, M., Gallucci, M., & Costantini, G. (2014). Safeguard power as a protection against imprecise power estimates. *Perspectives on Psychological Science*, *9*(3), 319–332. https://doi.org/10.1177/1745691614528519

Priester, J. R., & Petty, R. E. (1996). The gradual threshold model of ambivalence: Relating the positive and negative bases of attitudes to subjective ambivalence. *Journal of Personality and Social Psychology*, *71*(3), 431–449. https://doi.org/10.1037//0022-3514.71.3.431

Ridder, D. T. D. de, Lensvelt-Mulders, G., Finkenauer, C., Stok, F. M., & Baumeister, R. F. (2012). Taking stock of self-control: A meta-analysis of how trait self-control relates to a wide range of behaviors. *Personality and Social Psychology Review*, *16*(1), 76–99. https://doi.org/10.1177/1088868311418749

Ridderinkhof, K. R., Band, G. P. H., & Logan, G. D. (1999). A study of adaptive behavior: Effects of age and irrelevant information on the ability to inhibit one's actions. *Acta Psychologica*, *101*(2-3), 315–337. https://doi.org/10.1016/s0001-6918(99)00010-4

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Samrani, G., Marklund, P., Engström, L., Broman, D., & Persson, J. (2018). Behavioral facilitation and increased brain responses from a high interference working memory context. *Scientific Reports*, *8*(1), 1–14. https://doi.org/10.1038/s41598-018-33616-3

Spieler, D. H., Balota, D. A., & Faust, M. E. (2000). Levels of selective attention revealed through analyses of response time distributions. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 506–526. https://doi.org/10.1037/0096-1523.26.2.506

Stroop, J. R. (1935). Studies of interference in a letter naming task. *Bulletin of the Psychonomic Society*, *10*, 31–34.

Surrey, C., Dreisbach, G., & Fischer, R. (2017). Context-specific adjustment of cognitive control: Transfer of adaptive control sets. *Quarterly Journal of Experimental Psychology*, *70*(11), 2386–2401. https://doi.org/10.1080/17470218.2016.1239748

Topolinski, S., Bakhtiari, G., & Erle, T. M. (2016). Can I cut the Gordian tnok? The impact of pronounceability, actual solvability, and length on intuitive problem assessments of anagrams. *Cognition*, *146*, 439–452. https://doi.org/10.1016/j.cognition.2015.10.019

Torres-Quesada, M., Funes, M. J., & Lupiáñez, J. (2013). Dissociating proportion congruent and conflict adaptation effects in a Simon-Stroop procedure. *Acta Psychologica*, *142*(2), 203–210. https://doi.org/10.1016/j.actpsy.2012.11.015

Vadillo, M. A. (2019). Ego Depletion May Disappear by 2020. *Social Psychology*, *50*(5-6), 282–291. https://doi.org/10.1027/1864-9335/a000375

Vadillo, M. A., Gold, N., & Osman, M. (2018). Searching for the bottom of the ego well: Failure to uncover ego depletion in Many Labs 3. *Royal Society Open Science*, *5*(8). https://doi.org/10.1098/rsos.180390

Verbruggen, F. (2016). Executive Control of Actions Across Time and Space. *Current Directions in Psychological Science*, *25*(6), 399–404. https://doi.org/10.1177/0963721416659254

Vohs, K., Schmeichel, B. J., Lohmann, S., Gronau, Quentin, F., Finley, A. J., Ainsworth, S. E., Alquist, J. L., Baker, M. D., Brizi, A., Bunyi, A., Butschek, G. J., Campbell, C., Capaldi, J., Cau, C., Chambers, H., Chatzisarantis, N. L. D., Christensen, W. J., Clay, S. L., Curtis, J., … Albarracín, D. (2021). *A multi-site preregistered paradigmatic test of the ego depletion effect*. PsyArXiv. https://doi.org/10.31234/osf.io/e497p

Wenzel, M., Conner, T. S., & Kubiak, T. (2013). Understanding the limits of self-control: Positive affect moderates the impact of task switching on consecutive self-control performance. *European Journal of Social Psychology*, *43*(3), 175–184.

https://doi.org/10.1002/ejsp.1936

Wenzel, M., Kubiak, T., & Conner, T. S. (2014). Positive affect and self-control: Attention to self-control demands mediates the influence of positive affect on consecutive self-control. *Cognition and Emotion*, *28*(4), 747–755. https://doi.org/10.1080/02699931.2013.851069

Wilkowski, B. M., & Robinson, M. D. (2016). Cognitive Control Processes Underlying Individual Differences in Self-Control. In E. R. Hirt, J. J. Clarkson, & L. Jia (Eds.), *Self-regulation and ego control* (pp. 301–324). Academic Press. https://doi.org/10.1016/B978-0-12-801850-7.00015-9

Wright, R. A., Mlynski, C., & Carbajal, I. (2019). Outsiders' Thoughts on Generating Self-Regulatory-Depletion (Fatigue) Effects in Limited-Resource Experiments. *Perspectives on Psychological Science*, *14*(3), 469–480. https://doi.org/10.1177/1745691618815654

Xiao, S., Dang, J., Mao, L., & Liljedahl, S. (2014). When more depletion offsets the ego depletion effect. *Social Psychology*, *45*(5), 421–425. https://doi.org/10.1027/1864-9335/a000197

Xu, X., Demos, K. E., Leahey, T. M., Hart, C. N., Trautvetter, J., Coward, P., Middleton, K. R., & Wing, R. R. (2014). Failure to replicate depletion of self-control. *PLoS ONE*, *9*(10), e109950. https://doi.org/10.1371/journal.pone.0109950